ARTICLE   **OPEN**

Quantum tomography protocols with positivity are compressed sensing protocols

Amir Kalev[1], Robert L Kosut[2] and Ivan H Deutsch[1]

Characterising complex quantum systems is a vital task in quantum information science. Quantum tomography, the standard tool used for this purpose, uses a well-designed measurement record to reconstruct quantum states and processes. It is, however, notoriously inefficient. Recently, the classical signal reconstruction technique known as 'compressed sensing' has been ported to quantum information science to overcome this challenge: accurate tomography can be achieved with substantially fewer measurement settings, thereby greatly enhancing the efficiency of quantum tomography. Here we show that compressed sensing tomography of quantum systems is essentially guaranteed by a special property of quantum mechanics itself—that the mathematical objects that describe the system in quantum mechanics are matrices with non-negative eigenvalues. This result has an impact on the way quantum tomography is understood and implemented. In particular, it implies that the information obtained about a quantum system through compressed sensing methods exhibits a new sense of 'informational completeness.' This has important consequences on the efficiency of the data taking for quantum tomography, and enables us to construct informationally complete measurements that are robust to noise and modelling errors. Moreover, our result shows that one can expand the numerical tool-box used in quantum tomography and employ highly efficient algorithms developed to handle large dimensional matrices on a large dimensional Hilbert space. Although we mainly present our results in the context of quantum tomography, they apply to the general case of positive semidefinite matrix recovery.

INTRODUCTION

Determining an unknown signal from a set of measurements is a fundamental problem in science and engineering. However, as the number of free parameters defining the signal increases, its tomographic determination may become a daunting task. Fortunately, in many contexts there is prior information about the signal that may be useful for tomography. Compressed sensing[1–9] is a signal recovery technique developed for this aim. It utilises specific types of prior information about the structure of the signal to substantially compress the amount of information needed to reconstruct it with high accuracy. In particular, it harnesses the prior information that the signal has a concise representation, e.g., that it is a sparse vector with a few nonzero elements or a low-rank matrix with a few nonzero singular values. The compressed sensing protocol then defines special classes of measurements, henceforth referred to as 'compressed sensing measurements,' that enable the unique identification of the signal from within the restricted set of sparse vectors or low-rank matrices using substantially fewer measurement settings. Moreover, it provides algorithms for efficient reconstruction by defining a specific class of convex optimisation heuristics whose solution determine the unknown signal from the measurement outcomes with very high accuracy (see Methods). Importantly, solving any other optimisation programmes outside this class will not necessarily result in a compressed sensing protocol.

In the context of quantum information science, the 'signals' we seek to reconstruct are, for example, quantum states and processes, and the protocol for reconstruction is quantum tomography. Because the number of free parameters in quantum states and processes scale poorly (growing as some power of the total Hilbert space dimension, which in turn grows exponentially with the number of subsystems), there has been a concerted effort to develop techniques that minimise the resources necessary for tomography. To this end, the methodology of compressed sensing has been applied to the problem of quantum tomography.[10–21]

In the pioneering work of refs 11–13 it was proved that quantum measurements can be easily designed to be within the special class of measurements required for compressed sensing. Then, using the specifically chosen convex optimisation, low-rank density matrices (close to pure quantum states) or low-rank process matrices (close to unitary evolutions) can be accurately reconstructed with a substantially reduced number of measurement settings.

The work we report here identifies a critical link between quantum tomography and compressed sensing. We discuss in particular the case of quantum-state tomography, where the aim is to recover the density matrix, a positive semidefinite matrix, typically normalised with unit trace. We show that the positivity property alone imposes a powerful constraint that places strong restrictions on the physical states that are consistent with the data. As illustrated in Figure 1, this restriction is stronger than the one present in generic compressed sensing of signals that are not necessarily positive semidefinite matrices. This, in turn, has far reaching consequences. First and foremost, it implies that as long as quantum measurements are within the special class associated with compressed sensing, then any optimisation heuristic that contains the positivity constraint is effectively a compressed sensing protocol. Second, tools provided by the compressed sensing methodology now enable the construction of special

[1]Center for Quantum Information and Control, University of New Mexico, Albuquerque, NM, USA and [2]SC Solutions, Sunnyvale, CA, USA.
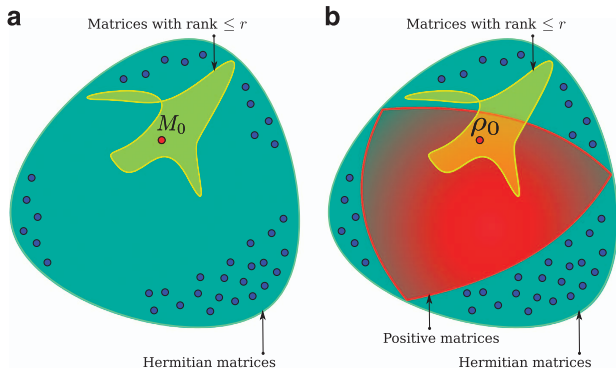Correspondence: A Kalev (amirk@unm.edu)

Figure 1. Schematic illustration of Theorem 1. (**a**) A generic compressed sensing scenario. The noiseless measurement record uniquely specifies the low-rank signal matrix M_0 (represented by a red dot) within the set of Hermitian matrices with rank $\leqslant r$ (the yellow non-convex set). However, there are many other Hermitian matrices with rank larger than r that are consistent with the measurement record (shown as the blue dots). (**b**) A generic scenario of compressed sensing of quantum states. If the noiseless measurement record comes from a density matrix, i.e., a positive matrix $\rho_0 \geqslant 0$ whose rank $\leqslant r$, then, according to Theorem 1, it specifies ρ_0 uniquely among the set of all positive matrices (shown as the red convex set). All other matrices that are consistent with the measurement record necessarily have negative eigenvalues and their rank is strictly larger than r.

types of informationally complete measurements that are robust to noise and to small model imperfections, with rigorous bounds. Finally, our results fundamentally unify many different quantum tomography protocols that were previously thought to be distinct, such as maximum-likelihood solvers, under the compressed sensing umbrella. We emphasise that constraining the normalisation (trace) to a fixed value, as one does for density matrices, has no role in the theorems we discuss below. Thus our results extend beyond the context of quantum-state tomography, applying, e.g., to process tomography when the latter is described by a completely positive map, and more generally to the reconstruction of low-rank positive semidefinite matrices.

RESULTS

Informational completeness

In quantum theory, a measurement is represented by a positive operator-valued measure, POVM, a set of positive semidefinite $d \times d$ matrices that form a resolution of the identity, $\mathcal{E} = \{E_\mu | E_\mu \geqslant 0, \sum_\mu E_\mu = 1\}$. The elements of a POVM represent the possible outcomes (events) of the measurement, and probability of measuring an outcome μ is given by the usual Born rule, $p_\mu = \mathrm{Tr}(E_\mu \rho)$, where ρ is the state of the system, a positive semidefinite matrix, $\rho \geqslant 0$, normalised such that $\mathrm{Tr}\rho = 1$. In the context of quantum-state tomography, informationally complete measurements have a central role. Let $\mathcal{S}$ be the set of all quantum states (density matrices). A measurement is said to be informationally complete if[22]

$$\forall \rho_a, \rho_b \in \mathcal{S}, \rho_a \neq \rho_b, \exists E_\mu \in \mathcal{E} \text{ s.t.} \mathrm{Tr}(E_\mu \rho_a) \neq \mathrm{Tr}(E_\mu \rho_b). \quad (1)$$

In other words, no two distinct states ρ_a and ρ_b yield the same measurement outcome probabilities. Thus, a (noise-free) record of an informationally complete measurement uniquely determines the state of the system. In general, for a d-dimensional Hilbert space, an informationally complete measurement consists of at least d^2 outcomes (POVM elements).

Although equation (1) gives a general definition of an informationally complete measurement, if one has prior

information about the state of the system, then we can make this definition more specific.[23,24] In particular, suppose the state is known *a priori* to be of a special class, $\mathcal{P}$, e.g., the class of density matrices of at most rank r. One defines a measurement to be $\mathcal{P}$ restricted informationally complete (restricted-IC) if it can only uniquely identify a quantum-state from within the subset $\mathcal{P}$, but cannot necessarily uniquely identify it from within the set of all quantum states. Such $\mathcal{P}$ restricted-IC measurements can be composed of fewer outcomes than the d^2 outcomes required for a general informationally complete measurement. For example, Heinosaari et al.[23] showed that when $\mathcal{P}$ is the set of density matrices of at most rank r, then rank-r restricted-IC measurements can be constructed with $\mathcal{O}(rd)$ outcomes, rather than $\mathcal{O}(d^2)$ outcomes required for a general informationally complete measurement. One can formalise this definition in the context of quantum-state tomography. A measurement is said to be $\mathcal{P}$ restricted-IC, if[23]

$$\forall \rho_a, \rho_b \in \mathcal{P}, \rho_a \neq \rho_b, \exists E_\mu \in \mathcal{E} \text{ s.t.} \mathrm{Tr}(E_\mu \rho_a) \neq \mathrm{Tr}(E_\mu \rho_b). \quad (2)$$

In some situations, a measurement can satisfy a stricter definition of informational completeness than the $\mathcal{P}$ restricted-IC of equation (2). A measurement is said to be $\mathcal{P}$ strictly-IC, if[24]

$$\forall \rho_a \in \mathcal{P}, \forall \rho_b \in \mathcal{S}, \rho_a \neq \rho_b, \exists E_\mu \in \mathcal{E} \text{ s.t.} \mathrm{Tr}(E_\mu \rho_a) \neq \mathrm{Tr}(E_\mu \rho_b). \quad (3)$$

There is a subtle yet important difference in the definitions of $\mathcal{P}$ restricted-IC and $\mathcal{P}$ strictly-IC. Whereas the measurement record of the former identifies a unique state within the set $\mathcal{P}$, the measurement record of a the latter identifies a unique state within the set of all quantum states. These notions of informationally completeness are key to understanding compressed sensing and its application in quantum tomography, as we discuss below.

The relation between informational completeness and compressed sensing

At its heart, the compressed sensing methodology employs prior information to reduce the number of measurements required to reconstruct an unknown signal. Here we consider the compressed sensing recovery of a $d \times d$ Hermitian matrix, M. Let the measurement record be specified as a vector-valued linear map, $y_i = \mathcal{A}[M]_i = \mathrm{Tr}(A_i M)$, where $\mathcal{A}$ is known as the 'sensing map.' In general, when the set $\{A_i\}$ forms a basis for $d \times d$ matrices with at least d^2 elements,[25] then the measurement record is informationally complete in the sense of equation (1), and in the absence of measurement noise, the signal can be recovered uniquely.

If, however, we know *a priori* that rank$(M) \leqslant r$, with $r \ll d$, then we can substantially reduce the number of measurement samples required to uniquely reconstruct the unknown signal matrix. This is codified in a theorem by Recht et al.[8] and Candès et al.[9] that we restate as follows:

Theorem (compressed sensing). Let the unknown signal M_0 be a Hermitian matrix with rank$(M_0) \leqslant r$, and let $\mathbf{y} = \mathcal{A}[M_0]$ be the measurement record obtained by a sensing map, $\mathcal{A}$, that corresponds to compressing sensing measurements for rank r. Then M_0 is the *unique* Hermitian matrix within the set of *low-rank* Hermitian matrices (up to rank r) that is consistent with the measurement record.

Importantly, in compressed sensing, when $r \ll d$, there are generally an infinite number of Hermitian matrices with rank larger than r that are consistent with the measurement record. Thus, the measurement record associated with compressed sensing cannot uniquely specify M_0 among all $d \times d$ Hermitian matrices, and therefore it is not informationally complete in the sense of equation (1). If, however, the sensing map $\mathcal{A}$ corresponds to compressed sensing measurements (e.g., it satisfies the restricted isometry property,[4] see Methods), then according to the above theorem, the measurement record uniquely specifies

M_0 within the restricted set of low-rank Hermitian matrices (rank $(M) \leqslant r \ll d$). Therefore compressed sensing measurements correspond to rank-r restricted-IC, in the sense of equation (2).

This relation between compressed sensing measurements and rank-r restricted-IC implies that any successful search must be restricted to the low-rank set of Hermitian matrices. To achieve this, one solves the convex optimisation problem,[8,9,21]

$$\hat{M} = \arg\min_{M} \|M\|_* \; \text{s.t.} \; \boldsymbol{y} = \mathcal{A}[M], \tag{4}$$

where $\|M\|_* = \mathrm{Tr}\sqrt{M^\dagger M}$, is the nuclear (or trace) norm, which serves as the convex proxy for rank minimisation. Under the conditions above, the optimal solution is $\hat{M} = M_0$, i.e., exact recovery. The use of the nuclear norm is essential here. If one uses only the compressed number of samples, then solving any other optimisation that is not related to the above rank-minimisation heuristic by some regularisation will not result in a successful recovery. For example, the solution of the convex programmes $\arg\min_{M} \mathrm{Tr}(M) \text{s.t.} \boldsymbol{y} = A[M]$, and $\arg\min_{M} \|\boldsymbol{y} - \mathcal{A}[M]\|_2$ with $m \ll d^2$ samples $\{y_i\}$ will generally yield a solution that is very different from M_0. Such estimators generally require $m \sim d^2$ samples to recover M_0. The analogous result holds for compressed sensing of sparse vectors. There ones require minimisation of the ℓ_1 norm of the vector, a convex heuristic for vector-sparsity.

In what follows, we specialise the compressed sensing paradigm to the case of positive matrix recovery, and particular to quantum-state tomography. There, the aim is to recover the state of the system, ρ, which has the key property of positivity, $\rho \geqslant 0$.

The role of positivity in compressed sensing quantum tomography

Our central result is summarised in the following theorem:

Theorem 1. Let P_0 be a positive semidefinite matrix with rank $(P_0) \leqslant r$, and let $\boldsymbol{y} = \mathcal{A}[P_0]$ be the measurement record obtained by a sensing map $\mathcal{A}$ that corresponds to compressing measurements for a rank-r Hermitian matrix. Then P_0 is the unique matrix within the set of positive semidefinite matrices of any rank that is consistent with the measurement record.

This is an analogous theorem to the one presented by Bruckstein et al.[26] for the case of positive sparse vector solutions for an underdetermined set of linear equations. Its proof as well as the details concerning the requirements on the sensing map are given in the Supplementary Information Section A. It also extends a result by Candès et al.[27] and Demanet and Hand[28] from rank-1 matrices to matrices with rank $\leqslant r$ for all permissible r.

Theorem 1 differs qualitatively from the standard compressed sensing theorem in a few key aspects. As discussed above, the general theory of compressed sensing guarantees that if the signal is a low-rank matrix with rank $\leqslant r$, and if the sensing map corresponds to compressed sensing measurements, then the measurement record uniquely specifies the unknown signal matrix, but only within the subset of matrices with rank $\leqslant r$. Theorem 1, on the other hand, states that if the matrix to be estimated is constrained to be a positive matrix (e.g., a density matrix), then the measurement record uniquely specifies the matrix from within the entire set of positive Hermitian matrices. Therefore, without the positivity constraint, compressed sensing measurements correspond to rank-r restricted-IC measurements of equation (2), whereas under positivity, the same measurements correspond to rank-r strictly-IC measurements of equation (3). This central result of Theorem 1 is illustrated in Figure 1.

The implication of Theorem 1 for quantum-state tomography is as follows. Suppose that the state of the system ρ_0, a positive semidefinite matrix, has rank $\leqslant r$. Assume that we have measured the system with a sensing map that satisfies the appropriate compressed sensing property, and obtained the (noiseless) measurement record $\mathcal{A}[\rho_0] = \boldsymbol{p}$. Then, according to Theorem 1, ρ_0 is the only density matrix within the set of positive Hermitian

matrices of any rank that yields the measurement probabilities $\boldsymbol{p}$. Geometrically, as observed in refs 27,28, Theorem 1 states that the rank-deficient subset of the positive matrices cone is 'pointed.' Therefore, under the promise that rank$(\rho_0) \leqslant r$ and $\mathcal{A}$ corresponds to compressed sensing measurements, the space of matrices ρ that satisfy $\mathcal{A}[\rho] = \boldsymbol{p}$ and the cone of positive matrices intersect in a single point $\rho = \rho_0$.

Theorem 1 implies that the solution set contains only one matrix, the density matrix ρ_0. It follows that we can use any optimisation method to search for it, and we are guaranteed to find it. Thus, we have the following result: Given a quantum measurement record $\boldsymbol{p} = \mathcal{A}[\rho_0]$, such that rank$(\rho_0) \leqslant r$, and where $\mathcal{A}$ corresponds to compressed sensing measurements, then the solution to

$$\hat{\rho} = \arg\min_{\rho} \mathcal{C}(\rho) \; \text{s.t.} \; \mathcal{A}[\rho] = \boldsymbol{p} \; \text{and} \; \rho \geqslant 0, \tag{5}$$

or to

$$\hat{\rho} = \arg\min_{\rho} \|\mathcal{A}[\rho] - \boldsymbol{p}\| \; \text{s.t.} \; \rho \geqslant 0, \tag{6}$$

where $\mathcal{C}(\rho)$ is a any convex function of ρ, and $\|\cdot\|$ is any norm function, is unique: $\hat{\rho} = \rho_0$. By confining the feasible set of matrices to positive matrices, we ensure that the measurement record uniquely identifies ρ_0 from the set of all density matrices, and thus any convex function of ρ or the measurement error may serve as a cost function. For example, this result applies to maximum-(log) likelihood estimation[29] where $\mathcal{C}(\rho) = -\log\left(\prod_{\mu} \mathrm{Tr}(E_\mu \rho)^{p_\mu}\right)$. We thus conclude that when the feasible set of density matrices is constrained to be physical (i.e., have positive eigenvalues), any quantum tomography protocol whose sensing map corresponds to compressed sensing measurements will exhibit the compressed sensing effect. We do not include a trace constraint in the convex programmes above. In the noiseless case considered here it is redundant. Because the data came from a trace-preserving quantum measurements, the unique solution must be a normalised quantum-state. As discussed in the Supplementary Information, the constraints $\rho \geqslant 0$ and $\mathrm{Tr}\rho = 1$, taken together, immediately imply that ρ_0 is the only density matrix consistent with the noiseless data. When we consider the important case of noisy measurements, the consequence trace constraint is nontrivial, as we discuss in the next section.

Robustness to measurement noise and model imperfection

So far, we have discussed the ideal case of a noiseless measurement record, where in the context of quantum tomography, $\boldsymbol{p}$ denoted a probability vector. The compressed sensing methodology, however, assures a robust reconstruction of the signal in the presence of measurement noise. Our analysis inherits this crucial feature. In a realistic scenario, we allow for a noisy measurement record, $\boldsymbol{f} = \mathcal{A}[\rho_0] + \boldsymbol{e}$, where we assume that the noise contribution can be bounded by some norm $\|\boldsymbol{e}\| \leqslant \varepsilon$. In the context of quantum tomography we consider $\boldsymbol{f}$ to denote a vector of the observed frequencies of measurement outcomes.

Theorem 1 ensures robust recovery of the positive density matrix if the noise level is small by solving any convex optimisation problem. Under the assumptions of Theorem 1, any convex minimisation problem that searches for a solution within the cone of positive matrices must yield a solution $\hat{\rho}$ such that $\|\hat{\rho} - \rho_0\| \leqslant g(\varepsilon)$, where $g(\varepsilon) \to 0$ as $\varepsilon \to 0$. From a geometrical point of view, when the noisy data arises from a rank-deficient state, as we gain data, there are fewer states that could have given rise to that data because the convex set of physical states is highly constrained near the point. In the idealised limit of noiseless data, there is only one state compatible with the data. Therefore, qualitatively, we expect a compressed sensing effect no matter how we search for the solution whenever the data arises from low-rank positive matrices. Quantitatively, of course, different

heuristics may perform differently, yielding different estimates. Choosing the best optimisation depends, in part, on the specific noise model. For example, in Supplementary Information Section B we derive a specific bound on the Frobenius (Hilbert–Schmidt) norm $\|\hat{\rho}-\rho_0\|_F = \sqrt{\mathrm{Tr}(\hat{\rho}-\rho_0)^2}$, where $\hat{\rho}$ is the solution of a non-negative least-squares program

$$\hat{\rho} = \arg\min_{\rho}\|\mathcal{A}[\rho]-\boldsymbol{f}\|_2 \text{ s.t. } \rho \geqslant 0. \qquad (7)$$

Whereas the normalisation constraint, $\mathrm{Tr}\rho = 1$ was unecessary in the noiseless case, in the case of a noisy measurement record, the convex optimisation is not guaranteed to obtain a normalised state. One can include the trace constraint in the optimisation, but it is generally unnecessary in the noisy case as well. In fact, sometimes one can actually improve the robustness to noise by choosing $\mathrm{Tr}\rho \neq 1$, as we discuss below. In general, the output of the optimisation should then be renormalized to give the final estimate.

We see this explicitly in ref. 11, where Gross et al. obtained a compressed sensing version of quantum-state tomography by solving the minimisation problem,

$$\min_{\rho}.\mathrm{Tr}\rho \text{ s.t.}\|\boldsymbol{f}-\mathcal{A}[\rho]\|_2 \leqslant \varepsilon, \rho \geqslant 0. \qquad (8)$$

This is equivalent to minimising the nuclear norm of ρ under the same constraints, i.e., when the feasible set is $\rho \geqslant 0$, then $\|\rho\|_* = \mathrm{Tr}\rho$. As noted above, minimising the trace of the matrix in the absence of the positivity constraint is not equivalent to minimising the nuclear norm, and therefore, would not achieve compressed sensing. Although both equation (7) and equation (8) are compressed sensing programmes, in general they return different estimations. However, in Supplementary Information Section C we show that the nonegative least-squares program,

$$\min_{\rho}.\|\mathcal{A}[\rho]-\boldsymbol{f}\|_2 \text{ s.t.}\mathrm{Tr}\rho = t,\ \rho \geqslant 0 \qquad (9)$$

is exactly equivalent to the nuclear-norm minimisation of equation (8) for a particular choice of t. This fact was observed empirically in a recent experiment by Smith et al.,[17] in which quantum-state tomography via continuous measurement was achieved at a equivalent rate by both least-squares and trace minimisation, with the positivity constraint included. The difference between the final estimate was attributed to a difference in the robustness of the two estimators to noise. As equations (8) and (9) are formally equivalent, the noisy measurement can be equivalently accommodated by solving (9) with a choice of t that depends on the noise bound ε. As always, we renormalize to obtain the final density matrix.

In addition to noise in the measurements, there can be imperfections in the model. When the sensing map satisfies the restricted isometry property, the compressed sensing methodology is not restricted to exact rank-deficient signal matrices. It also ensures the robust recovery of the dominant rank-r part of the density matrix. Our analysis shares this important and nontrivial property. Lemma 2 given in Supplementary Information Section B is the root of this feature.

We have shown that Theorem 1 implies that for a positive matrix recovery, compressed sensing measurements correspond to a stronger notion of informationally completeness—a rank-r strictly-IC. This implies that for quantum tomography we can construct robust measurements that are also rank-r strictly IC. The robustness to measurement noise and model imperfection is guaranteed by the compressed sensing theory. For example, in the context of a many-qubit system, Liu[12] showed that $\mathcal{O}(rd\mathrm{poly}(\log d))$ expectation values of Pauli products, $w = \otimes_{i=1}^{n}\sigma_{a_i}$, where $\sigma_a \in \{I, \sigma_x, \sigma_y, \sigma_z\}$, satisfy the restricted isometry property with overwhelming probability. Therefore, this set of expectation values is, with high probability, a robust rank-r strictly-IC measurement record. Similar results hold for sparse quantum process matrix reconstruction, e.g., it is shown in ref. 13 that if the sensing map is constructed from random input states, and random observables, then the restricted isometry holds with high probability.
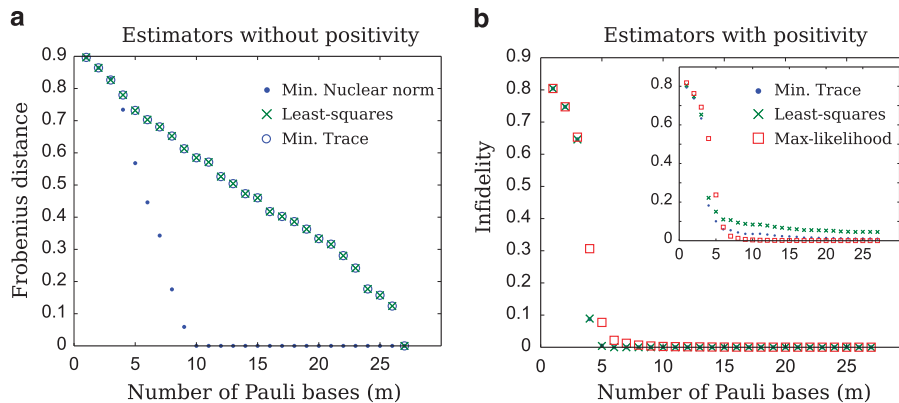

Figure 2. Comparison of different estimators with and without the positivity constraint. We simulate a three-qubit system in which we produce a pure state, $|\psi_0\rangle$, and perform random projective measurements in the Pauli basis (all simulations are averaged over 10 Haar-random states). (a) Estimation without the positivity constraint. We consider here the ideal case of a noiseless measurement record and plot the Frobenius distance between state $|\psi_0\rangle$ and the solution of an estimator, $\hat{\rho}$, $\sqrt{\mathrm{Tr}(\hat{\rho}-|\psi_0\rangle\langle\psi_0|)^2}$. The estimations are obtained by solving three different convex optimisations: (i) Nuclear-norm minimisation: equation (4), (ii) Least-squares minimisation: $\hat{\rho} = \arg\min_M\|\boldsymbol{p}-\mathcal{A}[M]\|_2$, and (iii) Trace minimisation: $\hat{\rho} = \arg\min_M\mathrm{Tr}(M)\text{s.t.}\boldsymbol{p} = \mathcal{A}[M]$. The figure clearly shows that only nuclear-norm minimisation achieves compressed sensing, i.e., exact recovery of the density matrix with a small number of measurement bases (here $m = 10$). Least-squares and trace minimisation require a full informationally complete measurement record with 27 Pauli bases to achieve exact recovery. (b) Estimation with the positivity constraint. We plot here the infidelity between $|\psi_0\rangle$ and the solution of different estimators, $1 - \langle\psi_0|\hat{\rho}|\psi_0\rangle$. The estimations are obtained by solving three different convex optimisations where the feasible set is constrained to the cone of positive matrices: (i) Nonnegative trace minimisation (equivalently nuclear-norm minimisation), equation (8) (ii) Nonnegative least-squares minimisation, equation (7), and (iii) The maximum-(log) likelihood estimator based on the algorithm described in ref. 30. In the main plot we simulate the case of an ideal noiseless measurement record; in the inset plot we simulate a statistically noisy measurement record that corresponds to frequency of outcomes for $N_{\mathrm{rep}} = 200$ repetitions. This figure exemplifies that when restricted to the set of positive matrices, all estimators are compressed sensing estimators.

Numerical test: compressed sensing state tomography of n-qubit system

Gross et al.[11,12,14] studied the problem of quantum-state tomography of an n-qubit system and showed that $m = \mathcal{O}(rd\,\mathrm{poly}(\log d))$ expectation values of random Pauli observables satisfy an appropriate restricted isometry property with high probability. If these expectation values are obtained through measurements in Pauli bases, i.e., local projective measurements on individual qubits in the eigenbasis of the Pauli observables, then, in fact, we obtain much more information. In addition to the average values, we also obtain the frequency of occurrence of each outcome, $E_\mu = \otimes_{i=1}^n P_{a_i}$, where μ indexes the series of a_i, $a = x, y, z$, and $P_{a_i} \in \{|\!\uparrow_{a_i}\rangle\langle\uparrow_{a_i}|, |\!\downarrow_{a_i}\rangle\langle\downarrow_{a_i}|\}$. Thus, we expect that we can obtain the required information for high-fidelity reconstruction using substantially fewer measurements based on individual outcomes in random Pauli bases rather than expectation values, and further reduce the resources needed for quantum-state tomography of a collection of qubits.

To exemplify this and the implication of Theorem 1, we perform numerical experiments on an n-qubit system (see Methods for details). In Figure 2, we simulate measurements on a three-qubit system, $d = 8$, and compare different numerical programmes to estimate the state. In Figure 2a we solve three estimators: equation (4) (nuclear-norm minimisation), $\min_M \|\boldsymbol{p} - \mathcal{A}[M]\|_2$ (least-squares), and $\min_M \mathrm{Tr}(M)\,\text{s.t.}\,\boldsymbol{p} = \mathcal{A}[M]$ (trace minimisation). Note that none of these estimators constrain the feasible set to the cone of positive matrices. The least-squares and trace minimisation are not convex heuristics for rank minimisation, and thus, as expected, they do not achieve compressed sensing. These programmes require a full informationally complete measurement record in order to reconstruct the quantum-state. On the other hand, the nuclear-norm heuristic does exhibit the compressed sensing effect, and recovers the density matrix with far fewer measurement outcomes. In Figure 2b we use the same data as in Figure 2a, but here we use estimators that restrict the feasible set to positive semidefinite matrices, e.g., the non-negative least-squares estimator, equation (7). The plots clearly show the implication of Theorem 1. Once restricted to the positive cone, the performance of all of the estimators is qualitatively the same and they all exhibit the compressed sensing effect. When the number of Pauli bases satisfy the appropriate restricted isometry property, the various estimators find the exact state in the idealised situation where the measurement record has no noise, and they find a good estimate that is close to the true state of the system in the presence of noise owing to finite sampling statistics of 200 repetitions.

In Figure 3, we treat a large dimensional Hilbert space: a ten qubit system, $d = 2^{10} = 1{,}024$. We simulate 30 random Pauli bases of a Haar-random pure state with $N_{\mathrm{rep}} = 100d$ repetitions for each observable. We estimate the state by solving equation (7) with a convex optimisation programme that can efficiently handle such large dimensional data sets (Riofrío, C. Personal communication). The programme implements a standard algorithm that uses gradient methods together with projection onto the positive cone. In the plot we see the compressed sensing effect due to the positivity constraint—all the information is captured in about 28 random Pauli bases, given sufficient statistics.

DISCUSSION

We have established a rigorous connection between the positivity property of quantum states and the compressed sensing method for quantum tomography. Thanks to the positivity constraint associated with physical states, the record of such measurements allows for a unique identification of a low-rank quantum-state within the set of all physical quantum states, of any rank. Thus, the measurements used for compressed sensing are informationally complete in a strict sense (strictly-IC). This aspect is fundamentally different than what
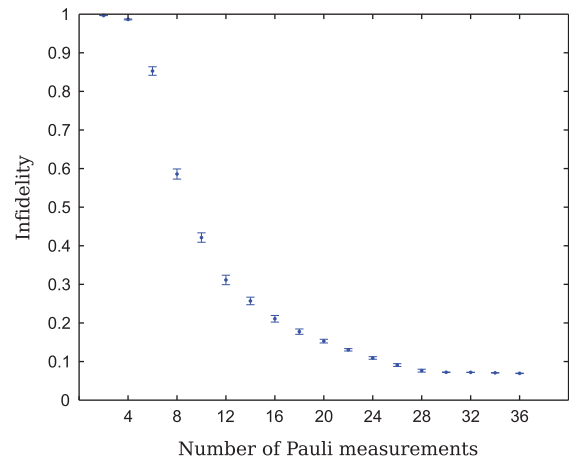

Figure 3. Ten-qubit state tomography. We simulate the data based on random Pauli-projective measurements (see text). The quantum tomography employs non-negative least-squares, according to equation (7). This algorithm can efficiently handle large dimensional matrices. We show the infidelity as a function of the number of measurement settings averaged over 10 Haar-random pure states (error bars shown). The simulation clearly exhibits the compressed sensing effect.

happens if positivity is not included. In the absence of the positivity constraint, the compressed sensing measurements are informationally complete in a restricted sense since they only uniquely identify a signal matrix from within the set of low-rank matrices (restricted-IC).

This strict relation has theoretical and practical implications. Most importantly, it implies that if one employs an optimisation programme that searches for a physical (positive) quantum-state, any quantum tomography procedure whose sensing map corresponds to compressed sensing measurements will exhibit the compressed sensing effect. This unifies apparently distinct numerical procedures such as maximum-likelihood and nuclear-norm minimisation under the umbrella of compressed sensing. From a practical perspective, when the positivity constraint is included, one can achieve compressed sensing estimation with any efficient convex optimisation, such as ADMM algorithms developed to handle large dimensional matrices. (A description of the ADMM algorithms and their application can be found at http://stanford.edu/boyd/admm.html.)

As the compressed sensing measurements that satisfy the restricted isometry property are robust to measurement noise and model imperfection, this allows us to construct strictly-IC measurements that are robust against such noise. That is, if there is measurement noise and/or we strictly violate the assumption that $\mathrm{rank}(\rho) \leqslant r$ (but only require that the density matrix is close to a density matrix with rank $\leqslant r$), then we are guaranteed that the estimation will be close to the unknown matrix.

Finally, though we have presented our results in the context of quantum-state tomography they are general and apply to the case of positive sparse vectors and positive rank-deficient matrices, the latter exemplified by quantum-process tomography.

METHODS

Compressed sensing measurements in matrix reconstruction

A sensing map for matrix reconstruction, $\mathcal{A}$, is defined as a vector-valued linear map on a $d \times d$ Hermitian matrix, $y_i = \mathcal{A}[M]_i$. This yields 'compressed sensing measurements for rank-r' if it guarantees a robust recovery of matrices with rank $\leqslant r$ by solving a nuclear-norm minimisation programme, e.g., the compressed sensing heuristic,

$$\hat{M} = \arg\min_M \|M\|_* \,\text{s.t.}\, \|\mathcal{A}[M] - \boldsymbol{f}\|_2 \leqslant \varepsilon, \tag{10}$$

where $\boldsymbol{f}$ is the noisy measurement record, $\boldsymbol{f} = \boldsymbol{y} + \boldsymbol{e}$. When the matrix is

promised to have rank r, the number of sufficient samples is of order $\mathcal{O}(rd)$, with possible logarithmic corrections, and the distance between the reconstruction $\hat{M}$ and M_0 is $\mathcal{O}(\varepsilon)$, where $\|e\|_2 \leqslant \varepsilon$. In this sense, the reconstruction is 'robust,' and compressed sensing when $r \ll d$. An analogous definition holds in the case of sensing maps for sparse vector reconstruction.

A sufficient condition that a sensing map yields compressed sensing measurements for matrix reconstruction is if it satisfies the 'restricted isometry property.' The map satisfies the restricted isometry property for rank-r if there is some constant $0 \leqslant \delta_r < 1$ such that,

$$(1-\delta_r)\|M\|_F^2 \leqslant \|\mathcal{A}[M]\|_2^2 \leqslant (1+\delta_r)\|M\|_F^2, \tag{11}$$

holds for all Hermitian matrices M with rank $\leqslant r$, where $\|M\|_F = \sqrt{\mathrm{Tr}(M^\dagger M)}$. The smallest constant δ_r for which this property holds is called the restricted isometry constant.

With small isometry constant δ_r, the sensing map $\mathcal{A}$ acts almost like an isometry when applied to rank $\leqslant r$ matrices, and thus allows us to effectively invert the measurement data to determine the matrix. Depending on the context, there are various results in the compressed sensing literature that apply for different values of the isometry constant. For example, Candés *et al.*,[9] show that the compressed sensing theory is applied when $\delta_{4r} < \sqrt{2} - 1$ (see Supplementary Information Section B).

Our results are general and apply whenever the sensing map corresponds to compressed sensing measurements that assures robust recovery through the solution of equation (10). Although the restricted isometry property is sufficient, our results are applicable in other cases, such as those described in ref. 21 where a robust recovery is guaranteed by $\mathcal{O}(rd)$ generic rank-one projections, or by $\mathcal{O}(rd\log(d))$ projectors onto random elements of an approximate four design.

Numerical experiments

In our numerical experiments, we simulate independent measurements of random Pauli bases on a Haar-random pure state of dimension $d = 2^n$, $\rho_0 = |\psi_0\rangle\langle\psi_0|$. The measurement record, given by the frequency of outcomes, $\boldsymbol{f}$, is generated by sampling N_{rep} times from the probability distribution $\boldsymbol{p} = \mathrm{Tr}(E\rho_0)$. Here E is the vector of POVM elements, each corresponding to a tensor product of projectors onto the eigenbasis of Pauli observables, $E_\mu = \otimes_{i=1}^n P_{a_i}$, where μ indexes the series of a_i, $a = x, y, z$, and $P_{a_i} \in \{|\uparrow_{a_i}\rangle\langle\uparrow_{a_i}|, |\downarrow_{a_i}\rangle\langle\downarrow_{a_i}|\}$. The measurement record is then used in various estimators. (Software for disciplined convex programming can be found at http://cvxr.com/.) We measure the performance by the average infidelity over 10 random pure states, $1 - \overline{\langle\psi_0|\hat{\rho}|\psi_0\rangle}$.

CONTRIBUTIONS

All authors contributed the ideas. AK performed the calculations. All authors wrote the manuscript.

COMPETING INTERESTS

The authors declare no conflict of interest.

REFERENCES

1. Donoho, D. L. Compressed sensing. *IEEE Trans. Inform. Theory* **52**, 1289–1306 (2006).
2. Candés, E. J., Romberg, J. K. & Tao, T. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inform. Theory* **52**, 489–509 (2006).
3. Candés, E. J., Romberg, J. K. & Tao, T. Stable signal recovery from incomplete and inaccurate measurements. *Commun. Pure Appl. Math.* **59**, 1207–1223 (2006).
4. Candés, E. J. The restricted isometry property and its implications for compressed sensing. *Comptes Rendus Mathematique* **346**, 589–592 (2008).
5. Candés, E. J. & Recht, B. Exact matrix completion via convex optimization. *Found. Comput. Math.* **9**, 717–772 (2009).
6. Candés, E. J. & Plan, Y. Matrix completion with noise. *Proc. IEEE* **98**, 925–936 (2010).
7. Candés, E. J. & Tao, T. The power of convex relaxation: Near-optimal matrix completion. *IEEE Trans. Inform. Theory* **56**, 2053–2080 (2010).
8. Recht, B., Fazel, M. & Parrilo, P. A. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM Rev.* **52**, 471–501 (2010).
9. Candés, E. J. & Plan, Y. Tight oracle inequalities for low-rank matrix recovery from a minimal number of noisy random measurements. *IEEE Trans. Inform. Theory* **57**, 2342–2359 (2011).
10. Kosut, R. L. Quantum Process Tomography via L1-norm Minimization. Preprint at http://arxiv.org/abs/0812.4323 (2008).
11. Gross, D., Liu, Y.-K., Flammia, S. T., Becker, S. & Eisert, J. Quantum State Tomography via Compressed Sensing. *Phys. Rev. Lett.* **105**, 150401 (2010).
12. Liu, Y.-K. Universal low-rank matrix recovery from Pauli measurements. *Adv. Neural Inform. Process. Syst.* **24**, 1638–1646 (2011).
13. Shabani, A. *et al.* Efficient measurement of quantum dynamics via compressive sensing. *Phys. Rev. Lett.* **106**, 100401 (2011).
14. Flammia, S. T., Gross, D., Liu, Y.-K. & Eisert, J. Quantum tomography via compressed sensing: error bounds, sample complexity and efficient estimators. *New J. Phys.* **14**, 095022 (2012).
15. Liu, W.-T., Zhang, T., Liu, J.-Y., Chen, P.-X. & Yuan, J.-M. Experimental quantum state tomography via compressed sampling. *Phys. Rev. Lett.* **108**, 170403 (2012).
16. Sanders, J. N. *et al.* Compressed sensing for multidimensional spectroscopy experiments. *J. Phys. Chem. Lett.* **3**, 2697–2702 (2012).
17. Smith, A. *et al.* Quantum control in the Cs $6S_{1/2}$ ground manifold using radio-frequency and microwave magnetic fields. *Phys. Rev. Lett.* **111**, 170502 (2013).
18. Schwemmer, C. *et al.* Experimental comparison of efficient tomography schemes for a six-qubit state. *Phys. Rev. Lett.* **113**, 040503 (2014).
19. Rodionov, A. V. *et al.* Compressed sensing quantum process tomography for superconducting quantum gates. *Phys. Rev. B* **90**, 144504 (2014).
20. Tonolini, F., Chan, S., Agnew, M., Lindsay, A. & Leach, J. Reconstructing high-dimensional two-photon entangled states via compressive sensing. *Sci. Rep.* **4**, 6542 (2014).
21. Kueng, R., Rauhut, H. & Terstiege, U. Low rank matrix recovery from rank one measurements. *Appl. Comput. Harmon. Anal.* (e-pub ahead of print 30 July 2015; doi: 10.1016/j.acha.2015.07.007).
22. Scott, A. J. Tight informationally complete quantum measurements. *J. Phys. A Math. Gen.* **39**, 13507 (2006).
23. Heinosaari, T., Mazzarella, L. & Wolf, M. M. Quantum Tomography under Prior Information. *Comm. Math. Phys.* **318**, 355–374 (2013).
24. Carmeli, C., Heinosaari, T., Schultz, J. & Toigo, A. Tasks and premises in quantum state determination. *J. Phys. A Math. Theory* **47**, 075302 (2014).
25. D'Ariano, G. M., Perinotti, P. & Sacchi, M. F. Informationally complete measurements and group representation. *J. Opt. B Quant. Semiclass. Opt.* **6**, S487 (2004).
26. Bruckstein, A. M., Elad, M. & Zibulevsky, M. On the uniqueness of nonnegative sparse solutions to underdetermined systems of equations. *IEEE Trans. Inform. Theory* **54**, 4813 (2008).
27. Candés, E. J., Strohmer, T. & Voroninski, V. PhaseLift: exact and stable signal recovery from magnitude measurements via convex programming. *Commun. Pure Appl. Math.* **66**, 1241–1274 (2013).
28. Demanet, L. & Hand, P. Stable optimizationless recovery from phaseless linear measurements. *J. Fourier Anal. Appl.* **20**, 199–221 (2014).
29. Hradil, Z. Quantum-state estimation. *Phys. Rev. A* **55**, R1561 (1997).
30. Teo, Y. S., Zhu, H., Englert, B.-G., Řeháček, J. & Hradil, Z. Quantum-state reconstruction by maximizing likelihood and entropy. *Phys. Rev. Lett.* **107**, 020404 (2011).

Supplementary Information accompanies the paper on the *npj Quantum Information* website (http://www.nature.com/npjqi)